

# **EXHIBIT B**

UNITED STATES DISTRICT COURT  
SOUTHERN DISTRICT OF NEW YORK

-----	X:	
THE NEW YORK TIMES COMPANY	:	
	:	
Plaintiff,	:	
	:	No. 1:23-CV-11195-SHS
-against-	:	
	:	
MICROSOFT CORPORATION, <i>et al.</i> ,	:	
	:	
Defendants.	:	
-----	X	

**DEFENDANT OPENAI OPCO, LLC'S SECOND SET OF REQUESTS  
FOR PRODUCTION OF DOCUMENTS AND THINGS (NOS. 62-104)**

In accordance with Federal Rules of Civil Procedure 26 and 34, Defendant OpenAI OpCo, L.L.C. (“OpenAI OpCo”), by and through its undersigned attorneys, hereby request that Plaintiff The New York Times Company responds to this Second Set of Requests for Production of Documents and Things (“Requests”) and produce for inspection and copying, within thirty (30) days of service, the Documents and Things described below, in accordance with the following Definitions and Instructions. OpenAI OpCo requests that Plaintiff produces Documents and other Things described below electronically or at the offices of Latham & Watkins LLP, 505 Montgomery Street, Suite 2000, San Francisco, California 94111.

### **DEFINITIONS**

1. Each of these Definitions and Instructions is incorporated into each of the Requests to which it pertains. Notwithstanding any Definition below, each word, term, or phrase used in these Requests is intended to have the broadest meaning permitted under the Federal Rules of Civil Procedure.

2. “Plaintiff,” “You,” and/or “Your” shall mean The New York Times Company, its officers, directors, managers, Employees, Agents, divisions, parent companies, subsidiaries, affiliates, including any Person who served in any such capacity at any time.

3. “Defendant” shall mean any defendant named in the Complaint filed on December 27, 2023 in *The New York Times Company v. Microsoft Corporation, et al.*, Case No. 1:23-cv-11195 and any subsequent complaint that You file in this action.

4. “OpenAI” shall mean OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI Opco L.L.C., OpenAI Global LLC, OAI Corporation, LLC, and OpenAI Holdings, LLC.

5. “Generative AI” shall mean models, algorithms, or other systems that are trained using data to generate new content (including text, images, or other media), such as ChatGPT,

Gemini, LLaMA, or Claude.

6. “Person” shall mean any natural person or any business, legal, or governmental entity or association.

7. “Communication” shall mean any transmission, exchange, or transfer of information by any means (in the form of facts, ideas, inquiries, or otherwise).

8. “Complaint” shall mean the Complaint filed on December 27, 2023 in *The New York Times Company v. Microsoft Corporation, et al.*, Case No. 1:23-cv-11195 and any subsequent complaint that You file in this action.

9. References to this “Action,” “Case,” “Litigation,” or “Lawsuit” mean the above-captioned matter, *The New York Times Company v. Microsoft Corporation, et al.*, Case No. 1:23-cv-11195, brought by Plaintiff pending in the United States District Court for the Southern District of New York.

10. “Agent” shall mean any Person asked, hired, retained, or contracted to assist You or act on Your behalf in connection with this Litigation whether employed or not by You.

11. “Employee” shall mean director, trustee, officer, employee, or partner of the designated entity, whether active or retired, full-time or part-time, current or former, and compensated or not.

12. “Document” shall mean all “writings” and “recordings” as those terms are defined in Federal Rules of Evidence 1001 and shall be afforded the broadest possible meaning permitted by Federal Rule of Civil Procedure 34. By way of illustration and without limitation, Documents include at least the following: originals, drafts and all non-identical copies of memoranda, reports, studies, surveys, analyses, books, manuscripts, notes, emails, graphs, notebooks, correspondence, interoffice communications, letters, diaries, calendars, photographs, motion pictures, sketches,

spreadsheets, source code, drawings, promotional material, technical papers, printed publications, and all other writings, as well as all non-paper information storage means such as sound reproductions, computer inputs and outputs, tape, film and computer memory devices, as well as tangible things such as models, modules, prototypes, and commercially saleable products. If a draft Document has been prepared in several copies that are not identical, or if the original identical copies are no longer identical due to subsequent notation, each non-identical Document is a separate Document. The foregoing specifically includes information stored in a computer database and capable of being generated in documentary form, such as electronic mail. For the avoidance of doubt, any electronic Documents or data shall include its associated metadata.

13. “Thing” shall have the full meaning ascribed to it by the Federal Rules of Civil Procedure.

14. “Asserted Work” shall mean any work listed in Exhibits A-I and K of the Complaint, or any additional work that may be listed on an amended complaint.

15. “Published Work” shall mean any work created by or published by You, including both Your Asserted Works and other works not referenced in the Complaint.

16. “ChatGPT” shall mean OpenAI's general purpose generative-AI model ChatGPT, and the underlying GPT-3.5 and GPT-4 large language models.

17. “GPT Services” shall mean OpenAI's general purpose generative-AI model ChatGPT, the underlying GPT-3.5 and GPT-4 large language models, and any GPT-based products, including any Application Programming Interface or other third-party services that rely on any of the aforementioned.

18. “Electronically Stored Information” includes, but is not limited to, e-mails and attachments, voice mail, instant messages, text messages, cell phone data and other electronic

communications, word processing documents, text files, hard drives, Excel spreadsheets and underlying formulae, graphics, audio and video files, databases, calendars, telephone logs, transaction logs, Internet usage files, offline storage or information stored on removable media (such as external hard drives, hard disks, floppy disks, memory sticks, flash drives, and backup tapes), information contained on laptops or other portable devices, and network access information and backup materials, TIF files, PDF files, Native Files and the corresponding metadata which is ordinarily maintained.

19. “Concern,” “Concerning,” “Related to,” and “Relating to,” are construed broadly to mean in whole or in part alluding to, responding to, concerning, relating to, connected with, involving, commenting on, in respect of, about, associated with, discussing, evidencing, showing, describing, reflecting, analyzing, summarizing, memorializing, consisting of, constituting, identifying, stating, tending to support, tending to discredit, referring to, or in any way touching upon.

20. “Exclusion Protocol” shall mean any rule, method, or policy to prevent specific user agents, such as web crawlers or bots, from accessing a website or parts of a website.

21. “NYT Annotated Corpus” shall mean the dataset You describe in the article “Fatten Up Your Corpus,” available at <https://open.nytimes.com/fatten-up-your-corpus-e72b22272222>.

22. “Including” means “including but not limited to.” “Include(s)” means “include(s) but is not limited to.”

23. OpenAI’s use of terms, phrases, and Definitions is for convenience and no term, phrase, or Definition shall be construed as an admission by OpenAI.

24. Any word that is not defined has its usual and customary meaning.

## **INSTRUCTIONS**

1. The following shall apply to all Requests:

- a. The singular of each word shall be construed to include its plural, and vice versa;
- b. The conjunctions “and” as well as “or” shall be construed either conjunctively or disjunctively, which is more inclusive in context;
- c. “Each” shall be construed to include “all” and “any,” and vice versa; and
- d. The present tense shall be construed to include the past tense, and vice versa.

2. If, in responding to these Requests, You encounter any ambiguities when construing a question, instruction, or definition, Your answer shall set forth the matter deemed ambiguous, and the construction used in answering.

3. In producing the Documents or Things requested, You are requested to produce all Documents and Things known or available to You regardless of whether such Documents and Things are possessed directly by You or Your Agents, Employees, representatives, investigators, or by Your attorneys or their Agents, Employees, representatives, or investigators.

4. In the event that any Document is withheld on a claim of attorney-client privilege or work-product doctrine on any other ground, You shall provide a log that identifies any such Document in a manner to be mutually agreed among the parties.

5. In producing the Documents and Things requested herein, please produce them in their original file folders, if any, or in lieu thereof, attach to the set of Documents produced from a given file or a photostatic or electronic duplicate of all written or printed material in the original file folder. In addition, the Documents shall be produced in the same sequence as they are contained or found in the original file folder. The integrity and internal sequence of the requested Documents within each file folder shall not be disturbed. Under no circumstances shall Documents

from any file folder be commingled with Documents from any other file folder.

6. If production of any requested Document(s) is objected to on the grounds that production is unduly burdensome, describe the burden or expense of the proposed discovery.

7. These Requests are continuing in nature. If You receive or otherwise become aware of information responsive to any Request after You have served Your responses to these Requests, You must promptly supplement Your responses to these Requests to provide such information, as required by Federal Rule of Civil Procedure 26.

### **REQUESTS FOR PRODUCTION**

#### **REQUEST FOR PRODUCTION NO. 62:**

All Documents and Communications relating to the provenance and process of publishing Your April 6, 2024 article titled “How Tech Giants Cut Corners to Harvest Data for A.I.”

#### **REQUEST FOR PRODUCTION NO. 63:**

All Documents and Communications relating to the provenance, production, and process of publishing Your April 16, 2024 The Daily podcast episode titled “A.I.’s Original Sin.”

#### **REQUEST FOR PRODUCTION NO. 64:**

All Documents and Communications relating to Your processes and discussions regarding the use and form of conflict-of-interest disclosures in Your Published Works.

#### **REQUEST FOR PRODUCTION NO. 65:**

All Documents and Communications relating to any efforts undertaken by You to enforce Your copyrights for any of Your Published Works against anyone, or to monitor for or investigate infringement of Your Published Works.

#### **REQUEST FOR PRODUCTION NO. 66:**

All Documents and Communications produced or provided to You by any third party in



this Litigation or concerning this Litigation.

**REQUEST FOR PRODUCTION NO. 67:**

All Documents and Communications regarding any requests by You to any third party requesting permission to quote, reference, or otherwise use in Your Published Works content not owned by You.

**REQUEST FOR PRODUCTION NO. 68:**

All Documents and Communications regarding licenses You have acquired in which You have obtained permission to quote, reference, or otherwise use in Your Published Works content not owned by You.

**REQUEST FOR PRODUCTION NO. 69:**

All Documents and Communications relating to any “benefits of [Your] work” that You were deprived of as a result of OpenAI’s conduct, as alleged in Paragraph 197 of the Complaint.

**REQUEST FOR PRODUCTION NO. 70:**

All Documents and Communications concerning (i) OpenAI’s alleged use of Your Published Works for training GPT Services; (ii) OpenAI’s relationship with Plaintiff; (iii) the impact of GPT Services on Your business and/or revenues; or (iv) the benefits of GPT Services to You or journalism generally.

**REQUEST FOR PRODUCTION NO. 71:**

All Documents and Communications relating to Your procedures and efforts to detect or prevent plagiarism in Your Published Works.

**REQUEST FOR PRODUCTION NO. 72:**

All Documents and Communications regarding projected or anticipated compensation for the exploitation of Your Published Works by any other person or entity, including but not limited

to, sales projections, royalty projections, future licensing plans, or any other Documents evidencing projections of future compensation.

**REQUEST FOR PRODUCTION NO. 73:**

All documents identified by You in Your Rule 26(a)(1) Initial Disclosures that are in Your possession, custody, or control.

**REQUEST FOR PRODUCTION NO. 74:**

All Documents pertaining to any allegations or claims raised by You relating to intellectual property infringement or any assertions of intellectual property rights, including Documents pertaining to proceedings involving such claims in any forum (including civil, criminal, and bankruptcy proceedings, as well as arbitrations and other quasi-judicial proceedings) to which You are or were a party or testifying witness, including, but not limited to, Documents sufficient to identify such proceedings; copies of any witness statements, declarations/affidavits, deposition transcripts, or other written or oral testimony You provided (sworn or unsworn); and the Documents pertaining to the final disposition of the matter, whether by judgment, settlement, dismissal, or other conclusion.

**REQUEST FOR PRODUCTION NO. 75:**

All Documents pertaining to any allegations or claims raised against You relating to intellectual property infringement or any assertions of intellectual property rights, including Documents pertaining to proceedings involving such claims in any forum (including civil, criminal, and bankruptcy proceedings, as well as arbitrations and other quasi-judicial proceedings), including, but not limited to, Documents sufficient to identify such proceedings; copies of any witness statements, declarations/affidavits, deposition transcripts, or other written or oral

testimony You provided (sworn or unsworn); and the Documents pertaining to the final disposition of the matter, whether by judgment, settlement, dismissal, or other conclusion.

**REQUEST FOR PRODUCTION NO. 76:**

All Documents and Communications regarding any study, survey, or other research regarding the market for Your Published Works or the impact of Generative AI on the journalism market, regardless of whether conducted by You or some other entity, from January 1, 2019 to present.

**REQUEST FOR PRODUCTION NO. 77:**

All Documents and Communication regarding Your allegation that OpenAI's conduct in this case was willful, as alleged in Paragraphs 124–126 and 165 of the Complaint.

**REQUEST FOR PRODUCTION NO. 78:**

All Documents sufficient to show all “copyright management information” associated with Your Published Works.

**REQUEST FOR PRODUCTION NO. 79:**

All Documents and Communications relating to Your coverage of lawsuits that You are a party to, including Your processes and discussions regarding the coverage of these lawsuits.

**REQUEST FOR PRODUCTION NO. 80:**

All Documents and Communications relating to formal employee evaluations for reporters who cover and report on lawsuits that You are a party to.

**REQUEST FOR PRODUCTION NO. 81:**

All Documents and Communications relating to employee evaluations outside Your formal review process for reporters who cover and report on lawsuits that You are a party to.

**REQUEST FOR PRODUCTION NO. 82:**

All Documents and Communications regarding Your application of “internal or third party . . . artificial intelligence large language models or applications . . . for the purposes of improving Times Services, for marketing research, to personalize marketing of Times Services to you and/or to create audiences for third party advertisers” as per Your Privacy Policy. See <https://help.nytimes.com/hc/en-us/articles/10940941449492-The-New-York-Times-Company-Privacy-Policy>.

**REQUEST FOR PRODUCTION NO. 83:**

All Documents and Communications concerning any Generative AI tools and product features, or prototypes thereof, You are building, considering implementing, and/or have implemented for Your use.

**REQUEST FOR PRODUCTION NO. 84:**

All Documents You considered for use as training data to train any Generative AI prototypes, tools, or products You are developing or have already implemented for Your use.

**REQUEST FOR PRODUCTION NO. 85:**

All Documents You used as training data to train any Generative AI prototypes, tools, or products You are developing or have already implemented for Your use.

**REQUEST FOR PRODUCTION NO. 86:**

All Documents and Communications concerning the rules or guidance for choosing what is and is not included in the training dataset for any Generative AI prototypes, tools, or products You are developing or have already implemented for Your use.

**REQUEST FOR PRODUCTION NO. 87:**

All Documents and Communications between You and any named plaintiff in *Daily News LP et al. v. Microsoft Corporation, et al.*, Case No 1:24-cv-03285 (S.D.N.Y.), directly or through

any Employee, Agent, or third party, related to copyright and artificial intelligence.

**REQUEST FOR PRODUCTION NO. 88:**

All Documents and Communications between You and the Center for Investigative Reporting, Inc. or its affiliated organizations (including but not limited to, Mother Jones, Reveal, and CIR Studios), directly or through any Employee, Agent, or third party, related to copyright and artificial intelligence.

**REQUEST FOR PRODUCTION NO. 89:**

All Documents and Communications between You and other news, media, or writers' organizations and publishers, as well as any other named plaintiff in *Tremblay v. OpenAI, Inc.*, No. 3:23-cv-03223 and *Silverman v. OpenAI, Inc.*, No. 3:23-cv-03416 in the Northern District of California, and *Authors Guild v. OpenAI Inc.*, No. 1:23-cv-08292, *Alter v. OpenAI Inc.*, No. 1:23-cv-10211, *Basbanes v. Microsoft*, No. 1:24-cv-00084, *The Intercept Media, Inc. v. OpenAI, Inc.*, No. 1:24-cv-01515, and *Raw Story Media, Inc. v. OpenAI, Inc.*, No. 1:24-cv-01514, *Daily News LP et al. v. Microsoft Corporation, et al.*, No. 1:24-cv-03285, *The Center for Investigative Reporting, Inc., v. OpenAI Inc.*, No. 1:24-cv-04872, in the Southern District of New York, directly or through any Employee, Agent, or third party, related to the use of published content for training Generative AI.

**REQUEST FOR PRODUCTION NO. 90:**

All Documents and Communications supporting Your contention that OpenAI competes with Your Published Works, as alleged in Paragraph 2 of the Complaint.

**REQUEST FOR PRODUCTION NO. 91:**

All Documents and Communications supporting Your contention that OpenAI "gave Times content particular emphasis when building [its] LLMs," as alleged in Paragraph 2 of the

Complaint.

**REQUEST FOR PRODUCTION NO. 92:**

All Documents and Communications regarding Your contention that OpenAI “create[s] products that substitute for The Times and steal audiences away from it,” as alleged in Paragraph 8 of the Complaint.

**REQUEST FOR PRODUCTION NO. 93:**

Documents and Communications sufficient to identify the harms You contend OpenAI has caused You by impeding Your ability to monetize Your content, as alleged in Paragraph 48 of the Complaint.

**REQUEST FOR PRODUCTION NO. 94:**

All Documents and Communications regarding Your contention that “synthetic search results divert important traffic away from copyright holders like The Times,” as alleged in Paragraph 110 of the Complaint.

**REQUEST FOR PRODUCTION NO. 95:**

Documents and Communications sufficient to reflect the revenue Wirecutter has generated via “affiliate referral,” as alleged in Paragraph 128 of the Complaint.

**REQUEST FOR PRODUCTION NO. 96:**

All Documents and Communications regarding any “decreases in traffic” or impact on the Wirecutter “advertising and subscription revenue” that You attribute to OpenAI or its products, as alleged in Paragraphs 128–29 of the Complaint.

**REQUEST FOR PRODUCTION NO. 97:**

All Documents and Communications regarding Your contention that “synthetic outputs” “erode consumer trust” in its products, such as Wirecutter, as alleged in Paragraph 133 of the

Complaint.

**REQUEST FOR PRODUCTION NO. 98:**

All Documents and Communications regarding Your contention that OpenAI and its products have “divert[ed] readers, including current and potential subscribers, away from The Times,” as alleged in Paragraph 157 of the Complaint.

**REQUEST FOR PRODUCTION NO. 99:**

All Documents and Communications regarding the development of The New York Times Annotated Corpus, including the decision to develop the Corpus.

**REQUEST FOR PRODUCTION NO. 100:**

All Documents and Communications regarding Your promotion and publicization of The New York Times Annotated Corpus.

**REQUEST FOR PRODUCTION NO. 101:**

All Documents and Communications regarding the decision to take down The New York Times Annotated Corpus.

**REQUEST FOR PRODUCTION NO. 102:**

All Documents and Communications regarding third parties’ access, download, and/or use of The New York Times Annotated Corpus.

**REQUEST FOR PRODUCTION NO. 103:**

All Documents and Communications regarding the Published Works that make up The New York Times Annotated Corpus.

**REQUEST FOR PRODUCTION NO. 104:**

All Documents and Communications regarding revenue received for use of The New York Times Annotated Corpus.

Dated: July 11, 2024

Respectfully Submitted

By: /s/ Joseph R. Wetzel

**LATHAM & WATKINS LLP**

Andrew M. Gass (*pro hac vice*)

*andrew.gass@lw.com*

Joseph R. Wetzel

*joseph.wetzel@lw.com*

505 Montgomery Street, Suite 2000

San Francisco, CA 94111

Telephone: 415.391.0600

Sarang V. Damle

*sy.damle@lw.com*

Elana Nightingale Dawson (*pro hac vice*)

*elana.nightingaledawson@lw.com*

555 Eleventh Street, NW, Suite 1000

Washington, D.C. 20004

Telephone: 202.637.2200

Allison L. Stillman

*alli.stillman@lw.com*

Luke A. Budiardjo

*luke.budiardjo@lw.com*

1271 Avenue of the Americas

New York, NY 10020

Telephone: 212.906.1200

By: /s/ Vera Ranieri

**MORRISON & FOERSTER LLP**

Joseph C. Gratz (*pro hac vice*)

*jgratz@mofo.com*

Vera Ranieri (*pro hac vice*)

*vrانieri@mofo.com*

425 Market Street

San Francisco, CA 94105

Telephone: 415.258.7522

Allyson R. Bennett (*pro hac vice*)

*abennett@mofo.com*

Rose S. Lee (*pro hac vice*)

*roselee@mofo.com*

Eric K. Nikolaides

*enikolaides@mofo.com*

707 Wilshire Boulevard, Suite 6000

Los Angeles, CA 90017-3543



Telephone: 213.892.5454

By: /s/ Katie Lynn Joyce

**KEKER, VAN NEST & PETERS LLP**

Robert A. Van Nest (*pro hac vice*)

*rvannest@keker.com*

Paven Malhotra (*pro hac vice*)

*pmalhotra@keker.com*

Michelle S. Ybarra (*pro hac vice*)

*mybarra@keker.com*

Nicholas S. Goldberg (*pro hac vice*)

*ngoldberg@keker.com*

Thomas E. Gorman (*pro hac vice*)

*tgorman@keker.com*

Katie Lynn Joyce (*pro hac vice*)

*kjoyce@keker.com*

Sarah Salomon (*pro hac vice*)

*ssalomon@keker.com*

R. James Slaughter (*pro hac vice*)

*rslaughter@keker.com*

633 Battery Street

San Francisco, CA 94111-1809

Telephone: 415.391.5400

*Attorneys for OpenAI Defendants*

**CERTIFICATE OF SERVICE**

I hereby certify that on July 11, 2024, a copy of the foregoing:

**DEFENDANT OPENAI OPCO, LLC'S SECOND SET OF REQUESTS FOR  
PRODUCTION OF DOCUMENTS AND THINGS (NOS. 62-104)**

was served by E-mail upon the following:

Genevieve Vose Wallace Ian B. Crosby Katherine Marie Peaslee SUSMAN GODFREY LLP 401 Union Street, Suite 3000 Seattle, WA 98101 Telephone: (206) 516-3880	Davida Brook Ellie Rae Dupler Emily K. Cronin SUSMAN GODFREY LLP 1900 Avenue of the Stars, Suite 1400 Los Angeles, CA 90067 Telephone: (310) 798-3150
Alexander Frawley Eudokia Spanos Tamar Lusztig Elisha Brandis Barron Zachary B. Savage SUSMAN GODFREY LLP One Manhattan West, 50 <sup>th</sup> Fl. New York, NY 10001 Telephone: (212) 336-8330	Jennifer Maisel Kristen Logan Steven M Lieberman ROTHWELL FIGG ERNST & MANBECK, P.C. 901 New York Avenue, N.W., Suite 900 East Washington, DC 20001 Telephone: (202) 783-6040
Scarlett Collings SUSMAN GODFREY LLP 1000 Louisiana Street, Suite 5100 Houston, TX 77002 Telephone: (713) 904-8812	
NYT-AI-SG-Service@simplelists.susmangodfrey.com	
Attorneys for New York Times Company	

Annette L. Hurst <b>ORRICK, HERRINGTON &amp; SUTCLIFF LLP</b> The Orrick Building 405 Howard Street San Francisco, CA 94105 Phone: 415.773.5700 Email: <i>ahurst@orrick.com</i> <i>djustice@orrick.com</i>	Christopher J. Cariello <b>ORRICK, HERRINGTON &amp; SUTCLIFF LLP</b> 51 West 2nd Street New York, NY 10019 Phone: 415.773.5700 Email: <i>ccariello@orrick.com</i>
Sheryl K. Garko Laura B. Najemy	Jeffrey S. Jacobson <b>FAEGRE DRINKER BIDDLE &amp; REATH</b> 1177 Avenue of the Americas

<b>ORRICK, HERRINGTON &amp; SUTCLIFF LLP</b> 222 Berkeley Street, Suite 2000 Boston, MA 02116 Phone: 617.880.1800 Email: <i>sgarko@orrick.com</i> <i>lnajemy@orrick.com</i>	New York, NY 10036 Phone: 212.248.3191 Email: <i>jeffrey.jacobson@faegredrinker.com</i>
Jared B. Briant <b>FAEGRE DRINKER BIDDLE &amp; REATH</b> 1144 Fifteenth Street, Suite 3400 Denver, CO 80202 Phone: 303.607.3500 Email: <i>jared.briant@faegredrinker.com</i>	
<i>nyclassactions_microsoft_ohs@orrick.com</i> <i>MicrosoftNYClassActionFDBR@faegredrinker.com</i>	
<i>Attorneys for Defendant Microsoft Corporation</i>	

Joseph R. Wetzel